

PCT

SAK

RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/647,019

TIME: 16:23:00

Input Set : A:\12525-407001.TXT

Output Set: N:\CRF3\07032001\I647019.raw

4 <110> APPLICANT: Harvey, Richard P.
5 Palmer, Stephen J.
6 Rosenthal, Nadia A.
7 Musaro, Antonio
9 <120> TITLE OF INVENTION: NOVEL MOLECULES EXPRESSED DURING MUSCLE
10 DEVELOPMENT AND GENETIC SEQUENCES ENCODING THE SAME
13 <130> FILE REFERENCE: 12525-407001
15 <140> CURRENT APPLICATION NUMBER: 09/647,019
C--> 16 <141> CURRENT FILING DATE: 2001-06-13
18 <150> PRIOR APPLICATION NUMBER: PCT/AU99/00220
19 <151> PRIOR FILING DATE: 1999-03-26
21 <150> PRIOR APPLICATION NUMBER: AU PP2634/98
22 <151> PRIOR FILING DATE: 1998-03-27
24 <160> NUMBER OF SEQ ID NOS: 20
26 <170> SOFTWARE: FastSEQ for Windows Version 4.0
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 778
30 <212> TYPE: DNA
31 <213> ORGANISM: Mus musculus
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (199)...(453)
37 <400> SEQUENCE: 1

| | |
|--|-----|
| 38 gctctcagga ctggagagag acagagcact ccagctatct cagccacatg aaaagcactg | 60 |
| 39 gaattgagat ccccgctcag aggacaccgg gagttccttc tctcctgtaa agcgcttttt | 120 |
| 40 gtgtttttgc acctggccgc ctgggactgt cctcaggcag taaaccaatc cagagagcag | 180 |
| 41 ggctaagacc ttgtgaat atg tcg aag cag cca att tcc aac gtc aga gcc | 231 |
| 42 Met Ser Lys Gln Pro Ile Ser Asn Val Arg Ala | |
| 43 1 5 10 | |
| 45 atc cag gcg aat atc aat att cca atg gga gcc ttt cgt ccg gga gct | 279 |
| 46 Ile Gln Ala Asn Ile Asn Ile Pro Met Gly Ala Phe Arg Pro Gly Ala | |
| 47 15 20 25 | |
| 49 ggg cag cct ccc aga agg aaa gag agt act cct gaa act gag gag gga | 327 |
| 50 Gly Gln Pro Pro Arg Arg Lys Glu Ser Thr Pro Glu Thr Glu Glu Gly | |
| 51 30 35 40 | |
| 53 gct cct acc acc tca gag gaa aag aag cca att cct gga atg aag aaa | 375 |
| 54 Ala Pro Thr Thr Ser Glu Glu Lys Lys Pro Ile Pro Gly Met Lys Lys | |
| 55 45 50 55 | |
| 57 ttt cca gga cct gtt gtc aac ttg tct gag atc caa aat gtt aaa agt | 423 |
| 58 Phe Pro Gly Pro Val Val Asn Leu Ser Glu Ile Gln Asn Val Lys Ser | |
| 59 60 65 70 75 | |
| 61 gaa ctg aaa ttt gtc ccc aaa ggt gaa cag tagtcgaaag gacacaaaag | 473 |
| 62 Glu Leu Lys Phe Val Pro Lys Gly Glu Gln | |
| 63 80 85 | |
| 65 ttacattgg atgcttagaa tcaggagatg catttcgttg acgtgttttt ccaagggaga | 533 |
| 66 aaaaacaatg ggttgaaata aacaacttcc tgaacatttt atacatttgt atgatgatca | 593 |
| 67 caaacctcct gaatgcccaa gactctagca aaaatatcct gtttgtacat ttatatttct | 653 |

ENTERED

RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/647,019

TIME: 16:23:00

Input Set : A:\12525-407001.TXT

Output Set: N:\CRF3\07032001\I647019.raw

```

68 tccttttact tggttgcatt tctcacttta gctacatttt tggcaccttg tagagcaaat 713
69 cagcacacga atttacaacc tgggaagtgt ggttttgagg agagatgtga tttttatgaa 773
70 gggggg 778
72 <210> SEQ ID NO: 2
73 <211> LENGTH: 85
74 <212> TYPE: PRT
75 <213> ORGANISM: Mus musculus
77 <400> SEQUENCE: 2
78 Met Ser Lys Gln Pro Ile Ser Asn Val Arg Ala Ile Gln Ala Asn Ile
79 1 5 10 15
80 Asn Ile Pro Met Gly Ala Phe Arg Pro Gly Ala Gly Gln Pro Pro Arg
81 20 25 30
82 Arg Lys Glu Ser Thr Pro Glu Thr Glu Glu Gly Ala Pro Thr Thr Ser
83 35 40 45
84 Glu Glu Lys Lys Pro Ile Pro Gly Met Lys Lys Phe Pro Gly Pro Val
85 50 55 60
86 Val Asn Leu Ser Glu Ile Gln Asn Val Lys Ser Glu Leu Lys Phe Val
87 65 70 75 80
88 Pro Lys Gly Glu Gln
89 85
91 <210> SEQ ID NO: 3
92 <211> LENGTH: 887
93 <212> TYPE: DNA
94 <213> ORGANISM: Homo sapiens
96 <220> FEATURE:
97 <221> NAME/KEY: CDS
98 <222> LOCATION: (185)...(448)
100 <400> SEQUENCE: 3
101 ggtttctcaat accgggagag gcacagagct atttcagcca catgaaaagc atcggaattg 60
102 agatcgagcgc tcagaggaca ccgggcgccc ctccacctt ccaaggagct ttgtattctt 120
103 gcattctggct gcctgggact tcccttaggc agtaaacaaa tacataaagc agggataaga 180
104 ctgc atg aat atg tcg aaa cag cca gtt tcc aat gtt aga gcc atc cag 229
105 Met Asn Met Ser Lys Gln Pro Val Ser Asn Val Arg Ala Ile Gln
106 1 5 10 15
108 gca aat atc aat att cca atg gga gcc ttt cgg cca gga gca ggt caa 277
109 Ala Asn Ile Asn Ile Pro Met Gly Ala Phe Arg Pro Gly Ala Gly Gln
110 20 25 30
112 ccc ccc aga aga aaa gaa tgt act cct gaa gtg gag gag ggt gtt cct 325
113 Pro Pro Arg Arg Lys Glu Cys Thr Pro Glu Val Glu Glu Gly Val Pro
114 35 40 45
116 ccc acc tcg gat gag gag aag aag cca att cca gga gcg aag aaa ctt 373
117 Pro Thr Ser Asp Glu Glu Lys Lys Pro Ile Pro Gly Ala Lys Lys Leu
118 50 55 60
120 cca gga cct gca gtc aat cta tcg gaa atc cag aat att aaa agt gaa 421
121 Pro Gly Pro Ala Val Asn Leu Ser Glu Ile Gln Asn Ile Lys Ser Glu
122 65 70 75
124 cta aaa tat gtc ccc aaa gct gaa cag tagtaggaag aaaaaaggat 468
125 Leu Lys Tyr Val Pro Lys Ala Glu Gln
126 80 85

```

RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/647,019

TIME: 16:23:00

Input Set : A:\12525-407001.TXT

Output Set: N:\CRF3\07032001\I647019.raw

```

128 tgatgtgaag aaataaagag gcagaagatg gattcaatag ctcactaaaa ttttatatat 528
129 ttgtatgatg attgtgaacc tcctgaatgc ctgagactct agcagaaatg gcctgtttgt 588
130 acatttatat ctcttccttc tagttggctg tatttcttac tttatcttca tttttggcac 648
131 ctcacagaac aaattagccc ataaattcaa cacctggagg gtgtgggtttt gaggagggat 708
132 atgattttat ggagaatgat atggcaatgt gcctaacgat tttgatgaaa agtttcccaa 768
133 gctacttcct acagtattttt ggtcaatatt tggaatgcgt tttagttctt caccttttaa 828
134 attatgtcac taaactttgt atgagttcaa ataaatattt gactaaatgt aaaatgtga 887

```

136 <210> SEQ ID NO: 4

137 <211> LENGTH: 88

138 <212> TYPE: PRT

139 <213> ORGANISM: Homo sapiens

141 <400> SEQUENCE: 4

```

142 Met Asn Met Ser Lys Gln Pro Val Ser Asn Val Arg Ala Ile Gln Ala
143 1 5 10 15
144 Asn Ile Asn Ile Pro Met Gly Ala Phe Arg Pro Gly Ala Gly Gln Pro
145 20 25 30
146 Pro Arg Arg Lys Glu Cys Thr Pro Glu Val Glu Glu Gly Val Pro Pro
147 35 40 45
148 Thr Ser Asp Glu Glu Lys Lys Pro Ile Pro Gly Ala Lys Lys Leu Pro
149 50 55 60
150 Gly Pro Ala Val Asn Leu Ser Glu Ile Gln Asn Ile Lys Ser Glu Leu
151 65 70 75 80
152 Lys Tyr Val Pro Lys Ala Glu Gln
153 85

```

155 <210> SEQ ID NO: 5

156 <211> LENGTH: 75

157 <212> TYPE: PRT

158 <213> ORGANISM: Xenopus laevis

160 <400> SEQUENCE: 5

```

161 Met Ser Lys Gln Pro Ala Ser Asn Ile Arg Ser Ile Gln Ala Asn Ile
162 1 5 10 15
163 Asn Ile Pro Met Gly Ala Phe Arg Pro Gly Ala Gly Gln Pro Pro Lys
164 20 25 30
165 Arg Lys Glu Phe Ser Thr Glu Glu Glu Gln His Val Pro Thr Pro Glu
166 35 40 45
167 Ser Glu Glu Lys Ser Glu Glu Lys Lys Pro Ile Pro Gly Ala Val Lys
168 50 55 60
169 Leu Pro Gly Pro Ala Phe Asn Leu Ser Glu Thr
170 65 70 75

```

172 <210> SEQ ID NO: 6

173 <211> LENGTH: 172

174 <212> TYPE: DNA

175 <213> ORGANISM: Homo sapiens

177 <400> SEQUENCE: 6

```

178 ggttctcaat accgggagag gcacagagct atttcagcca catgaaaagc atcggaattg 60
179 agatcgcagc tcagaggaca ccgggcgccc ctccacatt ccaaggagct ttgtattctt 120
180 gcatctggct gcctgggact tcccttaggc agtaaacaaa tacataaagc ag 172

```

182 <210> SEQ ID NO: 7

183 <211> LENGTH: 57

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/647,019

DATE: 07/05/2001

TIME: 16:23:00

Input Set : A:\12525-407001.TXT

Output Set: N:\CRF3\07032001\I647019.raw

```

184 <212> TYPE: DNA
185 <213> ORGANISM: Homo sapiens
187 <400> SEQUENCE: 7
188 ggataagact gcatgaatat gtcgaaacag ccagtttcca atgtagagc catccag      57
190 <210> SEQ ID NO: 8
191 <211> LENGTH: 87
192 <212> TYPE: DNA
193 <213> ORGANISM: Homo sapiens
195 <400> SEQUENCE: 8
196 gcaaatatca atattccaat gggagccttt cggccaggag cagggtcaacc cccagaaga      60
197 aaagaatgta ctccctgaagt ggaggag      87
199 <210> SEQ ID NO: 9
200 <211> LENGTH: 149
201 <212> TYPE: DNA
202 <213> ORGANISM: Homo sapiens
204 <400> SEQUENCE: 9
205 ggtgttcctc ccacctcgga tgaggagaag aagccaattc caggagcgaa gaaacttcca      60
206 ggacctgcag tcaatctatc ggaaatccag aatattaaaa gtgaactaaa atatgtcccc      120
207 aaagctgaac agtagtagga agaaaaaag      149
209 <210> SEQ ID NO: 10
210 <211> LENGTH: 422
211 <212> TYPE: DNA
212 <213> ORGANISM: Homo sapiens
214 <400> SEQUENCE: 10
215 gattgatgtg aagaaataaa gaggcagaag atggattcaa tagctcacta aaattttata      60
216 tatttgtatg atgattgtga acctcctgaa tgcctgagac tctagcagaa atggcctggt      120
217 tgtacattta tatctcttcc ttctagttgg ctgtatttct tactttatct tcatttttgg      180
218 cacctcacag aacaaattag cccataaatt caacacctgg aggggtgtggg tttagaggag      240
219 gatatgattt tatggagaat gatatggcaa tgtgcctaac gattttgatg aaaagtttcc      300
220 caagctactt cctacagtat tttggtcaat atttggaatg cgttttagtt cttcaccttt      360
221 taaattatgt cactaaactt tgtatgagtt caaataaata tttgactaaa tgtaaaatgt      420
222 ga      422
224 <210> SEQ ID NO: 11
225 <211> LENGTH: 40
226 <212> TYPE: PRT
227 <213> ORGANISM: Patinopecten sp.
229 <400> SEQUENCE: 11
230 Ser Val Ile Gln Arg Asn Ile Arg Lys Trp Val Leu Arg Leu Asn Trp
231 1 5 10 15
232 Gln Trp Trp Lys Leu Tyr Ser Lys Val Lys Pro Leu Leu Ser Ile Ala
233 20 25 30
234 Arg Gln Glu Glu Glu Met Lys Glu
235 35 40
237 <210> SEQ ID NO: 12
238 <211> LENGTH: 40
239 <212> TYPE: PRT
240 <213> ORGANISM: Rattus norvegicus
242 <400> SEQUENCE: 12
243 Leu Val Ile Gln Trp Asn Ile Arg Ala Phe Met Gly Val Lys Asn Trp

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/647,019

DATE: 07/05/2001

TIME: 16:23:00

Input Set : A:\12525-407001.TXT

Output Set: N:\CRF3\07032001\I647019.raw

```

244 1          5          10          15
245 Pro Trp Met Lys Leu Tyr Phe Lys Ile Lys Pro Leu Leu Lys Ser Ala
246          20          25          30
247 Glu Thr Glu Lys Glu Met Ala Asn
248          35          40
250 <210> SEQ ID NO: 13
251 <211> LENGTH: 40
252 <212> TYPE: PRT
253 <213> ORGANISM: Rattus norvegicus
255 <400> SEQUENCE: 13
256 Phe Cys Ile Gln Tyr Asn Ile Arg Ala Phe Met Asn Val Lys His Trp
257 1          5          10          15
258 Pro Trp Met Lys Leu Phe Phe Lys Ile Lys Pro Leu Leu Lys Ser Ala
259          20          25          30
260 Glu Thr Glu Lys Glu Met Ala Thr
261          35          40
263 <210> SEQ ID NO: 14
264 <211> LENGTH: 40
265 <212> TYPE: PRT
266 <213> ORGANISM: Homo sapiens
268 <400> SEQUENCE: 14
269 Glu Asn Asn Val Met Asn Ile Arg Gln Phe Asn Cys Ser Pro His Pro
270 1          5          10          15
271 Tyr Trp Leu Pro Asn Phe Met Asp Val Phe Thr Trp Ser Leu Pro Phe
272          20          25          30
273 Val Gly Glu Lys Arg Val Thr Glu
274          35          40
276 <210> SEQ ID NO: 15
277 <211> LENGTH: 6
278 <212> TYPE: PRT
279 <213> ORGANISM: Artificial Sequence
281 <220> FEATURE:
282 <223> OTHER INFORMATION: consensus sequence
284 <221> NAME/KEY: VARIANT
285 <222> LOCATION: (1)...(6)
286 <223> OTHER INFORMATION: Xaa = Any Amino Acid
288 <400> SEQUENCE: 15
W--> 289 Ile Gln Xaa Xaa Ile Arg
290 1          5
292 <210> SEQ ID NO: 16
293 <211> LENGTH: 42
294 <212> TYPE: PRT
295 <213> ORGANISM: Homo sapiens
297 <400> SEQUENCE: 16
298 Ser Glu Glu Asp Gly Phe Asp Gly Ala Thr Ala Ala Ala Arg Lys Glu
299 1          5          10          15
300 Val Ile Arg Trp Lys Ile Arg Ala Ile Gly Lys Met Ala Arg Val Phe
301          20          25          30
302 Ser Val Leu Arg Glu Glu Ser Glu Ser Val

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/647,019

DATE: 07/05/2001

TIME: 16:23:01

Input Set : A:\12525-407001.TXT

Output Set: N:\CRF3\07032001\I647019.raw

L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:289 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15